

AD-A092 171 RTE PRODUCTS CORP NEEDHAM HEIGHTS MA SYLVANIA SYSTEM--ETC F/G 17/2
WIDEBAND SPEECH MULTIPLE RATE PROCESSOR STUDY, PART I. (U)
OCT 79 R S CHEUNG, A J GOLDBERG DCA100-77-C-0054
UNCLASSIFIED NL

UNCLASSIFIED

END

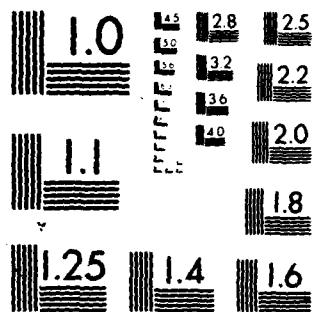
DATE

DATE
FILMED

81-52

DTIC

0710



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A092171

LEVEL

12
B.S.

FINAL REPORT

**WIDEBAND SPEECH
MULTIPLE RATE PROCESSOR STUDY**

DCA 100-77-C-0054

PART I

**DTIC
SELECTED
NOV 26 1980**

OCTOBER 1, 1979

GTE SYLVANIA
INCORPORATED
SYLVANIA SYSTEMS GROUP
EASTERN DIVISION

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

BDC FILE COPY

80 11 18 056

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
	AD-A092171	
4. TITLE (and Subtitle)	5. TYPE OF REPORT & PERIOD COVERED	
FINAL REPORT -	FINAL REPORT.	
Wideband Speech Multiple Rate Processor Study.	Sep 77 - Sep 79	
Part I.	6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s)	8. CONTRACT OR GRANT NUMBER(s)	
R. S. Cheung and A. J. Goldberg	DCA 100-77-C-0054	
9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	
GTE Sylvania Inc. ✓ GTE (Ind) 77 "A" Street Needham Heights, MA 02194		
11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE	
Defense Communications Agency Contract Management Division Code 200 Washington, D.C. 20305	October 4, 1979	
13. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	14. NUMBER OF PAGES	
	Part I - 47; Part II - 25	
	15. SECURITY CLASS. (of this report)	
	Unclassified	
	16a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report)		
Distribution of this document is unlimited. It may be released to the Clearinghouse, Department of Commerce for sale to the general public.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
Multiple rate processor, embedded coding schemes, non-embedded coding schemes, split-band voice coding, quadrature mirror filters.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)		
<p>→ This report presents the results of our investigations on the utility of multiple-rate processing (MRP) terminals in facilitating wideband/narrow-band communications. In addition, the design and development of a real-time embedded MRP scheme which transmits speech at the data rates of 2.4, 8.0, 9.6. and 16.0 Kb/s is discussed. →</p>		

(cont'd)

DD FORM 1 JAN 75 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

406451 *Em*

~~Unclassified~~

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract (Cont'd)

→ This report is written in two parts. Part I contains a description of non-embedded and embedded MRP schemes. Detailed treatments on two embedded algorithms, namely, the Linear Predictive Coding/Adaptive Predictive Coding with Adaptive Quantization (APC/APCQ); the Linear Predictive Coding/Split-Band Voice Coding (LPC/SBVC) are included. Part II contains the information on the real-time implementation of the LPC/SBVC coder on the government-owned Sylvania Programmable Signal Processors (PSP). The hardware and programming aspects of the high-speed multiplier-accumulator in the PSP's are also discussed.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

12

FINAL REPORT

WIDEBAND SPEECH MULTIPLE RATE PROCESSOR STUDY

DCA 100-77-C-0054

100-77-0054
RECEIVED
NOV 26 1980
C

PART I

October 1, 1979

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution unlimited

406451

TABLE OF CONTENTS

SECTION	TABLE OF CONTENTS	PAGE
	LIST OF ILLUSTRATIONS	iii
	LIST OF TABLES	iv
I	SUMMARY	1- 1
	1.1 Summary of the Program	1- 1
II	MULTIPLE-RATE SPEECH PROCESSING SYSTEMS	2- 1
	2.1 Introduction	2- 1
	2.2 Multiple-Rate Speech Processing Terminals	2- 4
	2.2.1 Non-embedded MRP Schemes	2- 5
	2.2.2 Embedded MRP Schemes	2-11
	2.3 The 2.4/16.0 Kb/s LPC/APCQ System	2-15
	2.4 The 2.4/8.0/9.6/16.0 Kb/s LPC/SBVC System	2-21
	2.4.1 The Theory of Quadrature Mirror Filter	2-29
	2.4.2 Results of the LPC/SBVC System	2-36
III	CONCLUSIONS AND RECOMMENDATIONS	3- 1
	3.1 Conclusions	3- 1
	3.2 Recommendations	3- 2
	REFERENCES	3- 3

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Avail and/or	
Dist	Special
A	

LIST OF ILLUSTRATIONS

FIGURE		PAGE
2- 1	Communications Between Narrowband and Wideband Terminals	2- 1
2- 2	Communications Between a Non-Embedded Multiple-Rate Processing and a Narrowband Terminal	2- 6
2- 3	Communications Between a Non-Embedded Multiple-Rate Processing and a Wideband Terminal	2- 7
2- 4	Communications Between Non-Embedded Multi-Rate Processing (MRP) Terminals	2- 9
2- 5	Communications Between Embedded Multi-Rate Processing Terminals	2-12
2- 6	Integrated 2400/16,000 BPS LPC/APCQ Voice Digitizer	2-16
2- 7	16,000 to 2400 BPS Tandem	2-19
2- 8	2400 to 16,000 BPS Tandem	2-20
2- 9	The LPC/SBVC MRP Transmitter	2-22
2-10	The LPC/SBVC MRP Receiver	2-23
2-11	2.4, 8.0, 9.6, 16.0 Kbps Tandem Configurations	2-26
2-12	1, 2, and 3-Bit Quantizers with Embedding Coding	2-28
2-13	Band-Splitting and Reconstruction Using Single Stage QMF	2-30
2-14	Magnitude Response of the 12-Tap QMF Filter	2-34
2-15	Composite Magnitude Response of a Split-Band Voice Coder (SBVC) Using 3 Stages of QMF Filters (No Quantization)	2-35

LIST OF TABLES

TABLE		PAGE
2-1	Input Bandwidth vs. Number of Residual Signal Quantizer Levels	2-15
2-2	Bit Allocations For 2.4/8.0/9.6/16.0 KB/s LPC/SBVC Scheme	2-25
2-3	Tabulation of Coefficients for Both the Low-Pass and High-Pass QMF Filters	2-33

SECTION I

Summary

1.1 Summary of the Program

In the DCA Wideband Speech Multiple-Rate Study (DCA 100-77-C-0054), GTE Sylvania investigated the utility of multiple-rate processing (MRP) terminals in wideband/narrowband communications. Particularly, two MRP schemes that employ embedded coding approaches were dealt with in detail. Also, GTE Sylvania developed a real-time simulation of the embedded Linear predictive Coder/Split Band Voice Coder which operates at the data rates of 2.4, 8.0, 9.6, and 16.0 Kb/s. This software was designed to run, in half-duplex mode, on the two GTE Sylvania Programmable Signal Processors (PSP) already owned by the Defense Communications Agency.

The DCA Wideband Speech Multiple-Rate Study was motivated by the unsatisfactory tandem performance obtained between a 2.4 Kb/s narrowband terminal (e.g., STU-2) and a 16 Kb/s wideband one (e.g., Tenley). The poor overall speech quality is generally attributed to the interface unit (e.g., Bellfield-Seeley Interface) which converts bit streams transmitted by one terminal into inputs of another. This conversion process includes synthesizing an estimate of the original waveform using one speech processing algorithm and then analyzing the resulting signal via another scheme. With this procedure, distortions are introduced in the processed speech which are due to the fact that most speech encoding methods are only optimized for clean input speech, but not for inputs that have been corrupted. Moreover, interactions between distortions of the first coder with that of the subsequent one further degrade the overall voice quality.

In light of the poor tandem performance between wideband and narrowband terminals, the objective of this study is to define ways that will improve it. Instead of devising a newer or better terminal interface, this effort investigates the alternative of replacing conventional wideband/narrowband terminals with multiple-rate processing (MRP) ones, each of which is capable of operating at several data

rates. In particular, this study highlights the embedded MRP schemes that employ only one speech processing algorithm for all transmissions. Two methods are dealt with in detail and they are:

1. a 2.4/16.0 Kb/s Linear Predictive Coding (LPC)/Adaptive Predictive Coding with Adaptive Quantization (APCQ)
2. a 2.4/8.0/9.6/16.0 Kb/s Linear Predictive Coding (LPC)/Split Band Voice Coding (SBVC)

The first algorithm employs conventional LPC for 2.4 Kb/s transmission, and by applying APCQ to the LPC residual, 16.0 Kb/s transmission results. Though the scheme produces highly intelligible speech at 2.4 Kb/s and good quality outputs at 16.0 Kb/s, reasonable performance cannot be attained at medium-band transmission (8-10 Kb/s) using LPC/APCQ owing to the tradeoffs between input bandwidth and levels of residual signal quantizers. In comparison, the second method, LPC/SBVC, is more versatile since it can operate at the data rates of 2.4, 8.0, 9.6, and 16.0 Kb/s. Similar to the first technique, LPC is also utilized for 2.4 Kb/s speech encoding. In this case, the LPC residual is split into eight subbands where each of them is individually quantized. Depending on the quantizer used, transmission at 8.0, 9.6, or 16.0 Kb/s is achieved. Though the LPC/SBVC operates at all rates of interest, unfortunately, quality obtained at the high rates is rather disappointing as compared to that of APCQ. This is attributed to the configuration of the quantizer in SBVC which is outside the prediction loop. Discussions of the above schemes can be found in Part I of this report.

Though the LPC/SBVC coder appears straightforward, its processing requirement is a combination of two coders, namely, LPC and SBVC. Henceforth, real-time implementation of the algorithm is generally unthinkable on many machines. To illustrate its complexity, LPC/SBVC transmitter functions include a LPC analyzer, computation of the LPC residual, three stages of split band filtering, and individual quantization of the eight subbands. Fortunately, the Sylvania PSP's, after modification under the Subband Coder Study (DCA 100-79-C-0001) are equipped with high speed multiplier-accumulators which can multiply two 16-bit numbers and accumulate the 32-bit product with 35 bit

precision in 206 nsec. Moreover, this hardware is especially efficient for linear filtering operations. With these PSP's, real-time implementations of the LPC/SBVC in half-duplex mode is possible. Flow charts and brief discussions of the real-time software are included in Part II of the report.

SECTION II

MULTIPLE-RATE SPEECH PROCESSING SYSTEMS

2.1 Introduction

Voice security terminals presently in use operate at different transmission rates, and in daily communications, subscribers of high data rate (wideband) terminals may have to converse with that of the low data rate (narrowband) ones. However, such a connection proves to be non-trivial owing to the fact that the terminals employ a variety of speech encoding schemes. One approach to facilitate such communication is to make use of a terminal interface which converts outputs of the first terminal into a new bit stream that is compatible with those of the subsequent terminal. As an example, the connection between a 2.4 Kb/s narrowband terminal which processes speech via Linear Predictive Coding (LPC) ¹ and a 16.0 Kb/s wideband terminal which employs Continuously Variable Slope Delta modulation (CVSD) ² is illustrated in Figure 2-1. To connect a call from the 2.4 Kb/s terminal to the 16.0 Kb/s one, synchronization has to be first established between the narrowband terminal and the interface and between the interface and the wideband terminal. Then the interface performs the decryption of the incoming LPC bit stream followed by the reconstruction of the input waveform using an LPC synthesizer. The resulting waveform is processed by a CVSD analyzer, and the 16 Kb/s output after encryption are transmitted to the wideband terminal. Though this connection seems transparent to the users, synchronization has to be maintained between the two subscribers and the interface throughout the entire conversation. Also shown in Figure 2-1, the terminals are secure (black), that is, only encrypted data are generated. However, "clear" data and speech (in digital or analog form) are created within the interface in such a way that communications between terminals are protected only if the entire interface is secure. This restriction increases the cost of the interfaces which may consequently limit their availability to the subscribers.

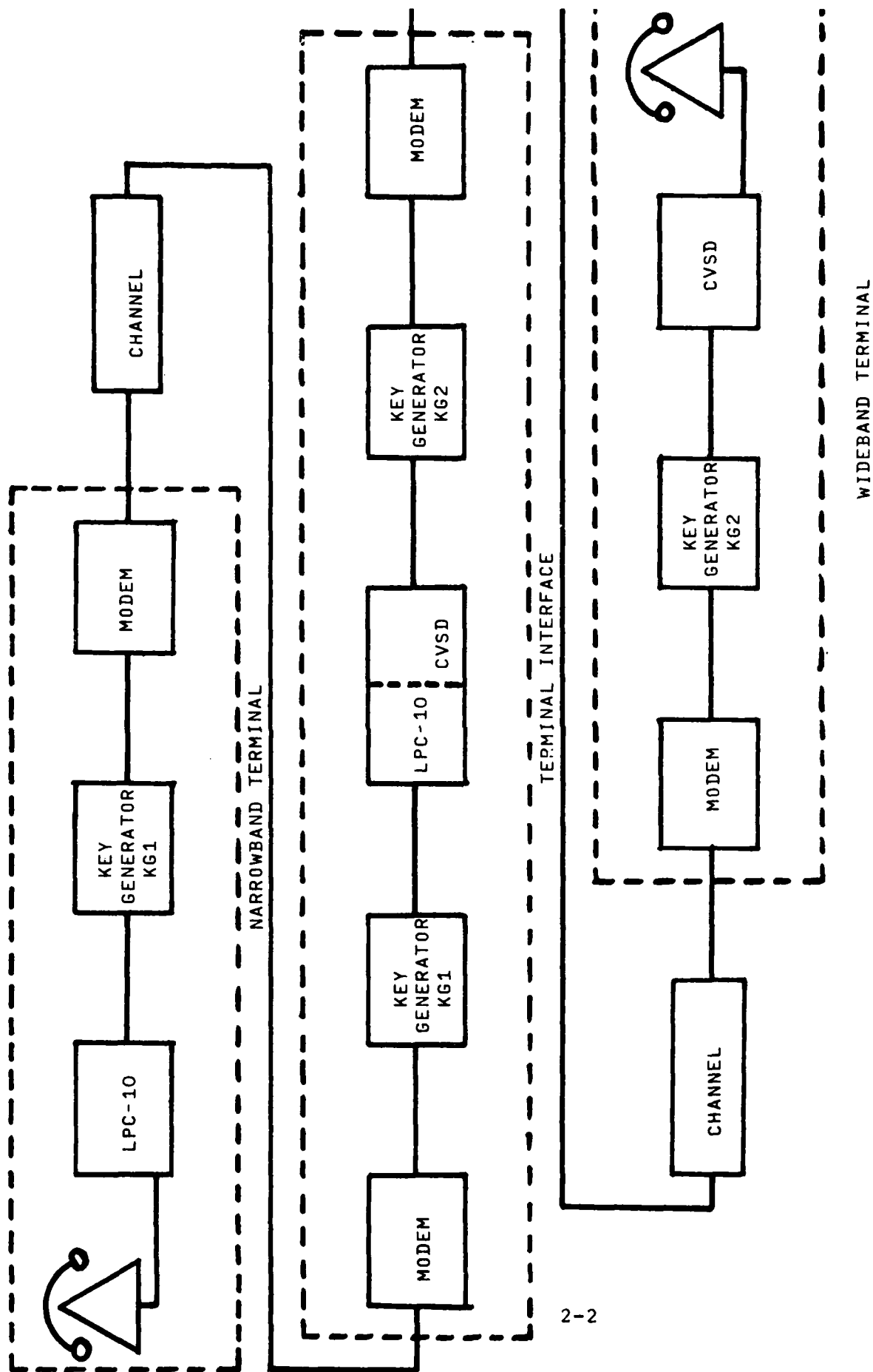


FIGURE 2-1 COMMUNICATIONS BETWEEN NARROWBAND AND WIDEBAND TERMINALS

Furthermore, though terminals may individually produce acceptable outputs, the overall speech quality obtained when they are connected through an interface is less than satisfactory. This is exemplified by the more muffled and buzzy speech quality obtained when a 16 Kb/s CVSD terminal is hooked up to a 2.4 Kb/s LPC terminal.³ The distortions can be attributed to the fact that the speech processing algorithms are only optimized for clean speech, but not for inputs derived from another speech encoding scheme. Moreover, these degradations are further compounded by the interactions of distortions introduced by the first speech processing method with that of the subsequent one.

In light of the costs and speech distortions associated with the terminal interfaces, it is desirable to find other means that will improve the tandem performance of the terminals. The obvious solution is to use only one type of terminal for all transmissions. In other words, identical terminals that operate at one data rate are utilized, henceforth, communications between subscribers require no interface hardware. Unfortunately, differences in transmission media tremendously limit the effectiveness of such communication systems. Another approach to improve the tandem performance between wideband and narrowband terminals is to develop new or optimize existing speech processing algorithms in the terminals in such a way that they will yield good quality outputs regardless of the input material. However, these techniques may not exist or be readily available and this approach may not be possible. A more viable method is to employ terminals known as multiple-rate processing (MRP) terminals each of which has a (or a collection of) speech encoding algorithm(s) that is capable of transmitting at various data rates. So, instead of the fact that users of either narrowband or wideband terminals can only function at a fixed rate, MRP terminal users have a choice of several transmission modes. As long as both subscribers are set on the same data rate, communications between them can be established without any terminal interfaces and this eliminates the tandeming problem.

2.2 Multiple-Rate Speech Processing Terminals

A multiple-rate processing (MRP) terminal is basically one that utilizes a single voice processor for both wideband and narrowband speech transmission. The heart of the processor is a speech processing scheme that is capable of encoding speech at a list of data rates. In general, there are two types of MRP schemes, namely, embedded and non-embedded methods. The non-embedded MRP algorithm utilizes a combination of several independent speech processing techniques that operate at different rates. In this situation, subscribers of a terminal can select from the available rates one which is compatible with that of the other terminal. Since the coder's algorithms are independent of each other and as a result, they can be individually optimized with respect to channel characteristics, speech quality, computation, and hardware requirements. Unfortunately, these MRP terminals requires a complicated and sometimes cumbersome call-up procedure in order to set up identical rates for both terminals. Another MRP strategy is to employ an embedded coding technique where the binary bit stream of the lower data rate scheme is buried in the outputs of the higher rate. So in contrast to non-embedded schemes where speech processing algorithms utilized are independent of each other, the lower data rate scheme of the embedded MRP terminal is a subset of the higher rate. Moreover, to facilitate communications between wideband and narrowband terminal users, an intelligent switch has to be inserted between the embedded MRP terminals in order to perform the stripping or filling in of the bits required for the particular rate of transmission. When compared to the tandem interface as shown in Figure 2-1, this switch does not convert the bit stream into analog waveform and then resample the waveform. Consequently, all problems associated with the interface are not present in these intelligent switches. However, constraints have to be incorporated in the design of higher data rate scheme, and this significantly increases the complexity of the embedded MRP scheme. The following sections discuss in detail the operations of both types of MRP terminals and their utility in narrowband/wideband communications.

2.2.1 Non-embedded MRP Schemes

The non-embedded MRP scheme is basically a collection of several independent speech processing algorithms. Employing either a manual or an automatic procedure, the subscribers can select one of the available schemes for transmitting or receiving speech. In general, the lower data rate technique yields intelligible speech of synthetic quality whereas the higher rate scheme produces a more natural sounding output. An example of the above MRP terminal is the bi-rate STU-2 terminal which has a 2.4 Kb/s Linear Predictive Coder (LPC) and a 9.6 Kb/s Adaptive Predictive Coder (APC).⁴ Since the two coders within the terminal bear no relationship to each other, they can function independently. The lower rate LPC is known to yield speech with a "buzzy" and unnatural quality whereas the higher rate APC results in slightly granular but more pleasing processed speech.

To detail the utility of the non-embedded MRP scheme, a block diagram depicting the connection of an MRP and a narrowband terminal is shown in Figure 2-2. The MRP terminal shown is a tri-rate terminal which can transmit at 2.4, 9.6, and 16.0 Kb/s, whereas the narrowband terminal shown functions only at 2.4 Kb/s. So, in order for the MRP terminal to communicate with the narrowband one, it is clear that the speech processing schemes at 2.4 Kb/s for both terminals have to be identical. Hence, if LPC is utilized in the narrowband terminal, the 2.4 Kbps coder of the MRP terminal also has to be an LPC. To establish a call between the two terminals, protocol information concerning the data rates has to be passed between them. In practice, initiating a call from the MRP terminal may require the subscriber to set manually the mode of the terminal to be 2.4 Kbps. On the other hand, if users of the narrowband terminals wish to speak to those of the MRP ones, information about the 2.4 Kbps data rate has to be first transmitted to the MRP terminal. From it, the mode of the MRP terminals can be automatically switched to 2.4 Kbps. Then communications between the terminals can commence. Utilizing a similar procedure, communications between a MRP and a wideband terminal that operates at 16 Kb/s, as shown in Figure 2-3, can also be established.

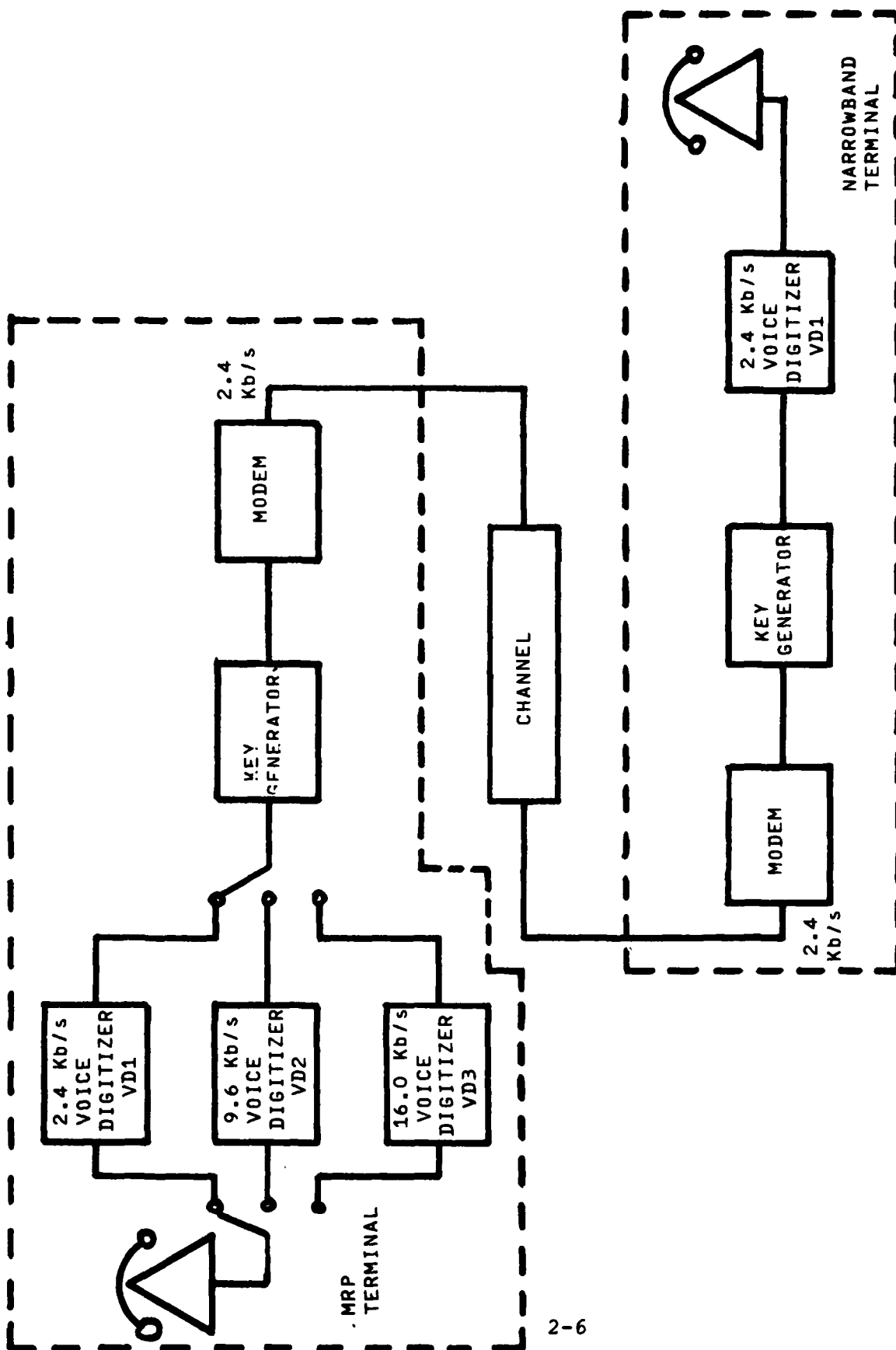
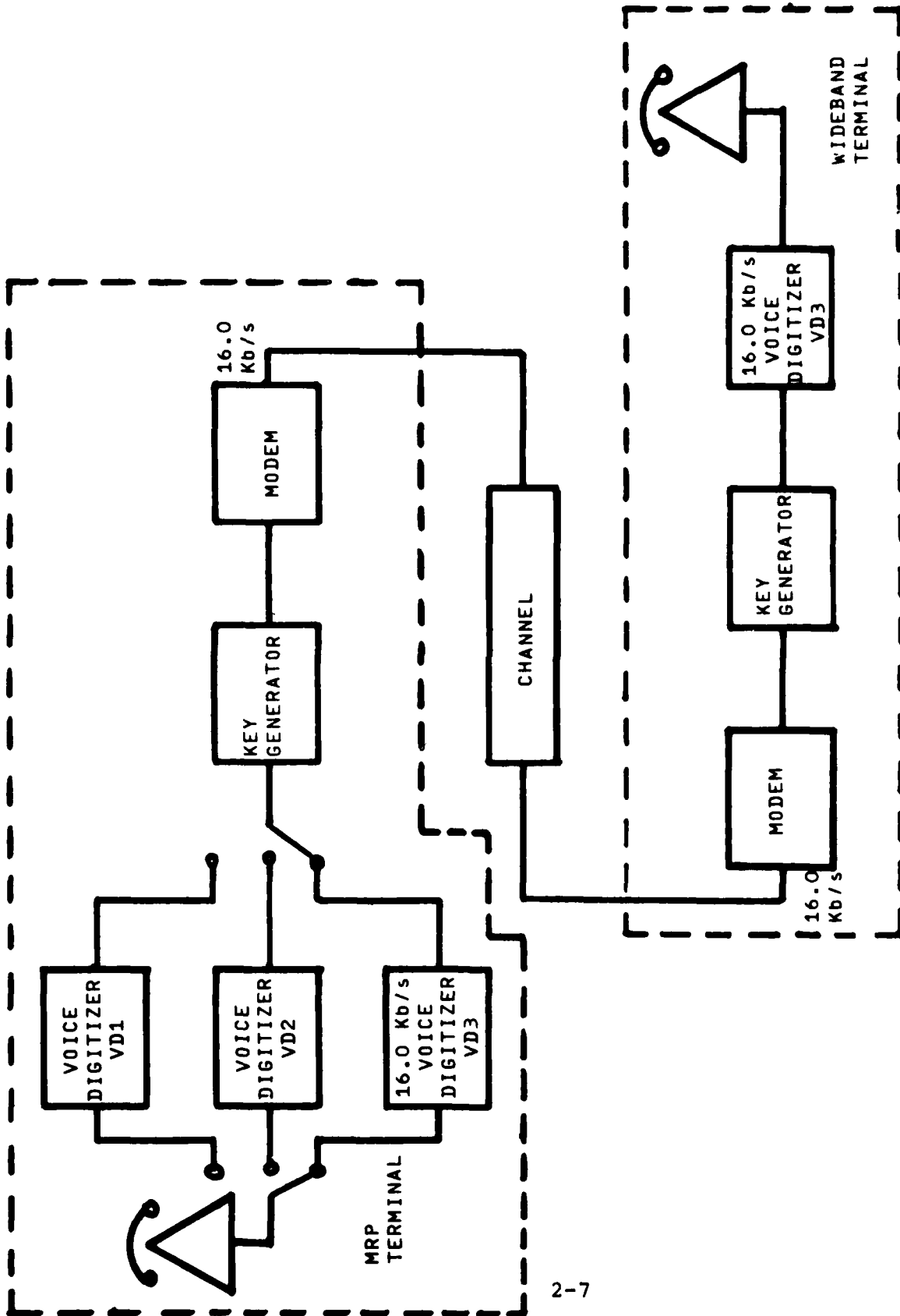


FIGURE 2-2 COMMUNICATIONS BETWEEN A NON-EMBEDDED MULTIPLE RATE PROCESSING AND A NARROWBAND TERMINAL



2-7

FIGURE 2-3 COMMUNICATIONS BETWEEN A NON-EMBEDDED MULTI-PROCESSING AND A WIDEBAND TERMINAL

Though connections between MRP and narrowband or MRP and wideband terminals are straightforward, communications between two non-embedded MRP terminals are rather complicated owing to the variety of available rates. Ideally, subscribers of these terminals may wish to transmit and receive at the highest possible rate for the purpose of achieving the best voice quality. Unfortunately, this may not always be possible due to the limitations of the channels. Furthermore, these subscribers may not necessarily want to restrict their utility of MRP terminals only to the lowest rate that only results in speech of vocoder quality. So, instead of fixing the communications between MRP terminals to a pre-determined and fixed data rate, an ideal alternative is that the MRP terminals can change and adjust their data rates according to media/threat conditions of the channel. To illustrate the above idea, a block diagram depicting the connection of two non-embedded MRP terminals is shown in Figure 2-4. Initially, these terminals may start up at the highest data rate and synchronization between the units is attempted with the transmission of preambles. Depending on the channel conditions, synchronization may fail between the units. Then the systems will automatically switch to a lower data rate mode and synchronization procedures are repeated. This process is iterated until reliable synchronization can be established. Throughout the entire conversation, the data rate at which synchronization is achieved has to be employed.

It is clear that the utilization of the non-embedded MRP terminals eliminates the wideband/narrowband tandeming problem since these terminals can function at different data rates. Another advantage is that the MRP terminals can be used in the existing media, such as commercial telephone networks and no additional modification or hardware (e.g., tandem interface) is required in the networks. One further point is that the speech processing algorithms within each telephone terminal can be optimized with respect to each individual method since they are independent of each other.

However, an apparent disadvantage of the above MRP terminal is that in order to setup the same data rate between the terminals,

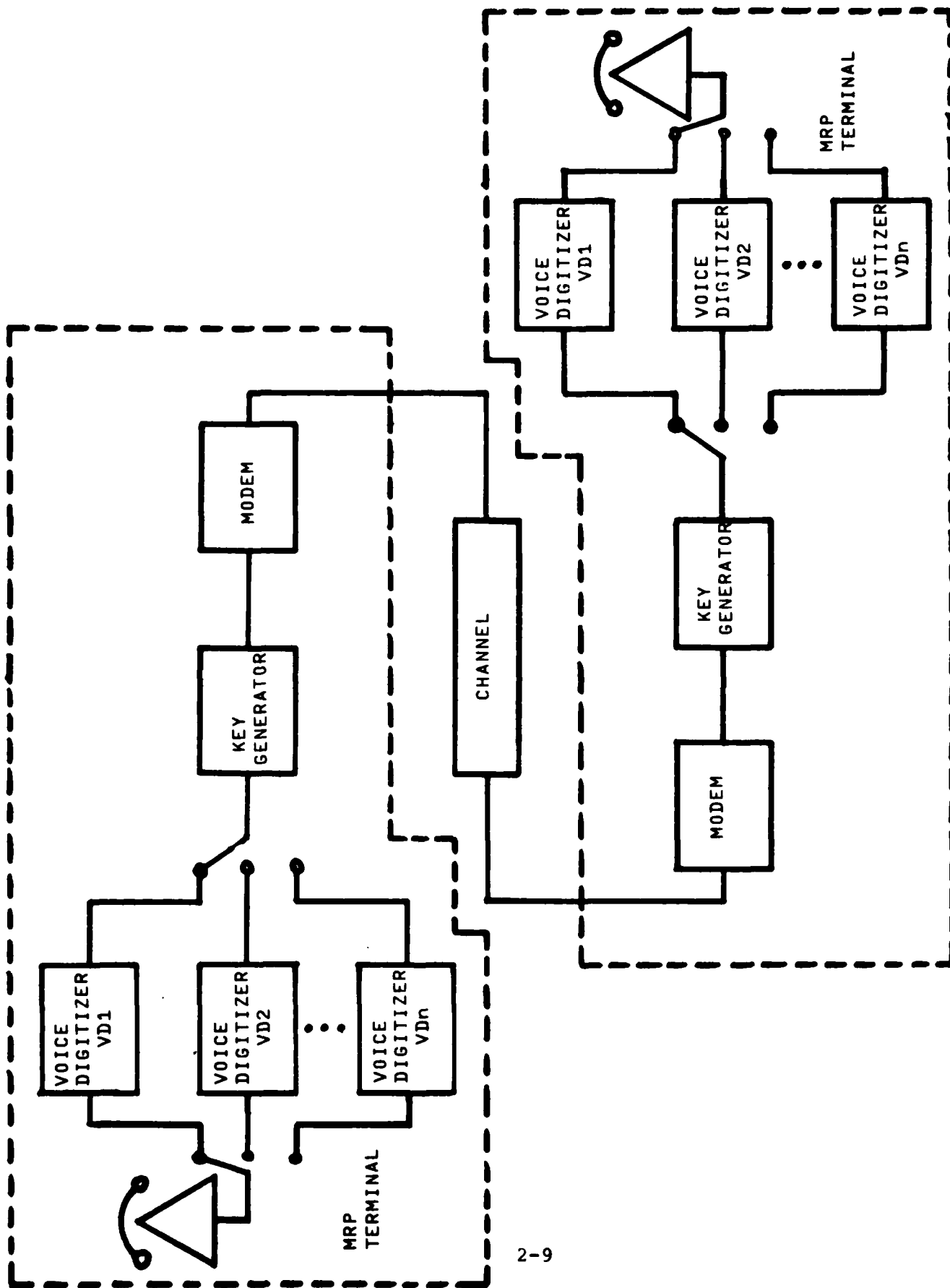


FIGURE 2-4 COMMUNICATIONS BETWEEN NON-EMBEDDED MULTI-RATE PROCESSING (MRP) TERMINALS

a complicated calling procedure has to be setup. This may include the polling of the terminals (MRP or independent ones) which determines the correct data rate needed to facilitate an efficient transmission. Also, the nonuniformity of the available speech processing algorithms at the different data rates can tremendously increase the complexity of the terminal hardware.

2.2.2 Embedded MRP Schemes

A relatively new concept to facilitate wideband/narrowband communications without the tandeming problem is to make use of MRP terminals that employ embedded coding schemes. In this approach, a single processing algorithm encodes speech at different data rates by allocating various number of bits to quantize the transmission parameters. Hence, quantizers with more levels are used to convert these variables into bit streams for the higher data rate transmission, and, consequently, this results in processed speech of good quality. On the other hand, quantizers with fewer bits can be applied to the same set of transmission parameters, and the resulting lower data rate method yields processed speech of poorer quality. Since in both rates the same speech processing algorithm is utilized and the same transmission variables are computed, the only difference between them is on the quantization of the outputs. So, by properly designing these quantizers, it is conceivable that bit streams needed for synthesizing the lower data rate schemes can be derived from that of the higher one, and these class of speech processing methods are, in general, referred to as embedded coding schemes.

To illustrate the utility of this algorithm in wideband/narrowband communication, connection between two such MRP terminals is shown in Figure 2-5. In this figure, the bottom MRP terminal is assumed to be functioning at 16 Kb/s while the top terminal is transmitting at 2.4 Kb/s. To complete this connection, a switch has to be utilized to perform the data rate conversion. In contrast to conventional switches whose sole function is to route calls from one location to another, the switch, as shown in Figure 2-5, is intelligent in the sense that it performs bit strippings or insertions. In addition, functions such as the determination of the two data rates, synchronizations between the switch and the MRP terminals, have to be done by this switch. For instance, if the user of the 16 Kb/s terminal wishes to converse with that of the 2.4 Kb/s one, connections between the 16 Kb/s terminal and the intelligent switch have to be established initially. Then protocol infor-

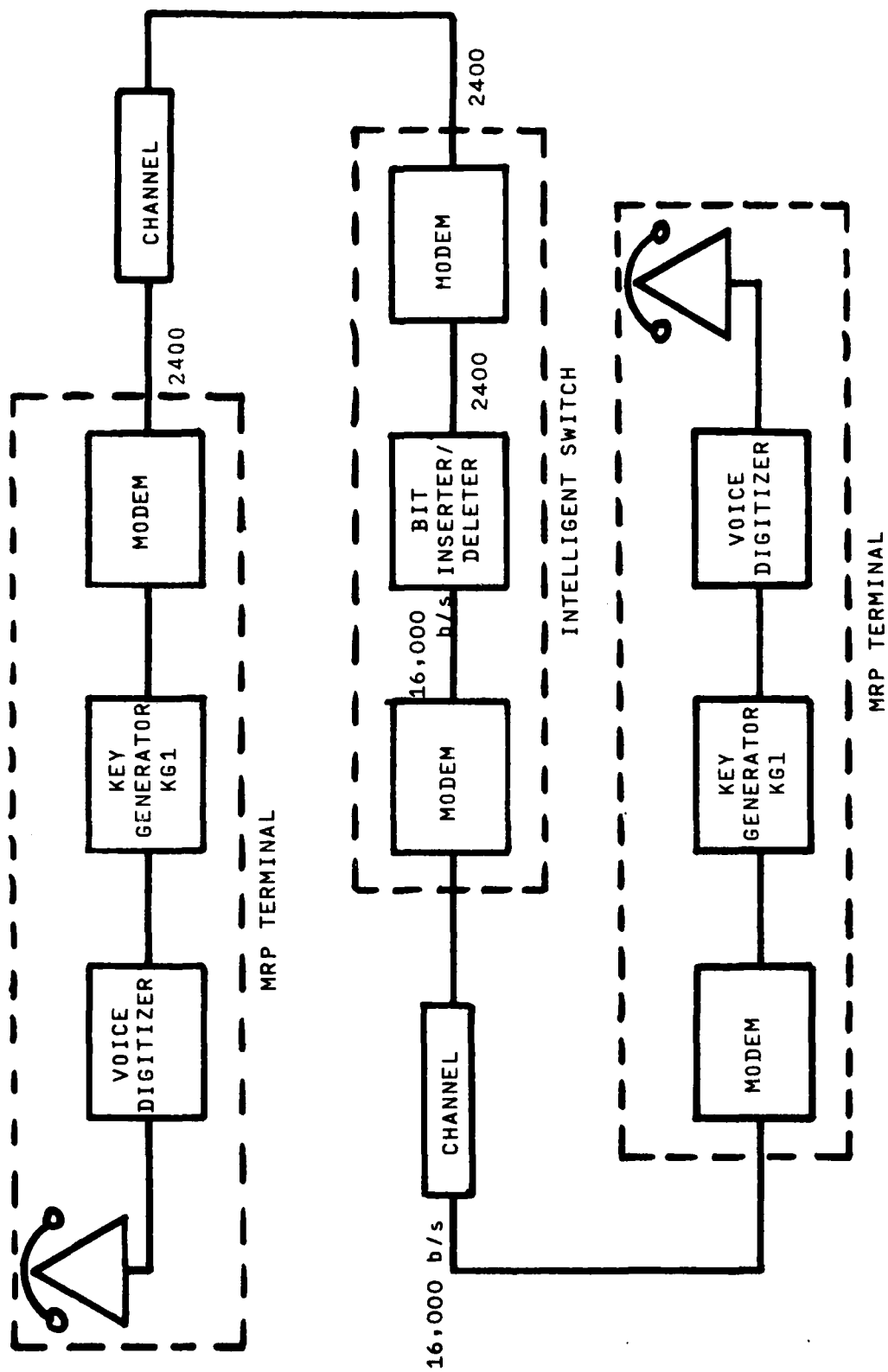


FIGURE 2-5 COMMUNICATIONS BETWEEN EMBEDDED MULTI-RATE PROCESSING TERMINALS

mation pertinent to the caller's terminal is transmitted to the switch via a preamble. From this, synchronization between the 16 Kb/s terminal and the switch is made, and, in a similar fashion, synchronization between the switch and the 2.4 Kb/s terminal is also established. After the synchronization process, transmission of data can commence, and the switch will be responsible for stripping from the 16 Kb/s data stream the bits needed to synthesize 2.4 Kb/s and transmitting the resulting bits to the 2.4 Kb/s terminal.

Unfortunately for this scheme, the design for the higher data rate is more complicated than it is necessary since bit streams for the high rate have to include those of the lower rates. Sometimes this results in inefficient coding of the transmission parameters.

However, there are a lot of advantages associated with this embedded scheme. As shown in the Figure 2-5, no tandem interface is needed as far as the need to resynthesize the waveform and then redigitize again. Consequently, the operations involved in the intelligent switch are also secure (black) and this simplifies the overall security requirements. Not only does this "embedded" multirate processing (MRP) approach eliminate tandeming requirements, it permits designers to consider a number of interesting options when transmitting voice over packet switching networks. For example, when a packet switched voice network is lightly loaded, the subscribers can send at 16,000 bps. As the network becomes saturated, however, "data reducers" located within the switching system can reduce the 16,000 bps transmissions to 2400 bps transmissions merely by stripping off the extra 13,600 bps. Voice quality drops, of course, but extra voice channel capacity is obtained and the network can now accept more subscribers. Thus, during light loading, the subscribers receive good service in the form of high voice quality but during peak loading, which under conventional system design would produce unacceptably long delays, users can still communicate, but with degraded voice quality.

In this report, two examples of embedded MRP schemes are dealt with in detail and they are:

1

i) a 2.4/16.0 Kbps Linear Predictive Coding (LPC)/Adaptive
Predictive Coding with Adaptive Quanti-
zation (APCQ)

ii) a 2.4/8.0/9.6/16.0 Kbps Linear Predictive Coding (LPC)/
Split Band Voice Coding (SBVC)

2.3 The 2.4/16.0 Kb/s LPC/APCQ System⁵

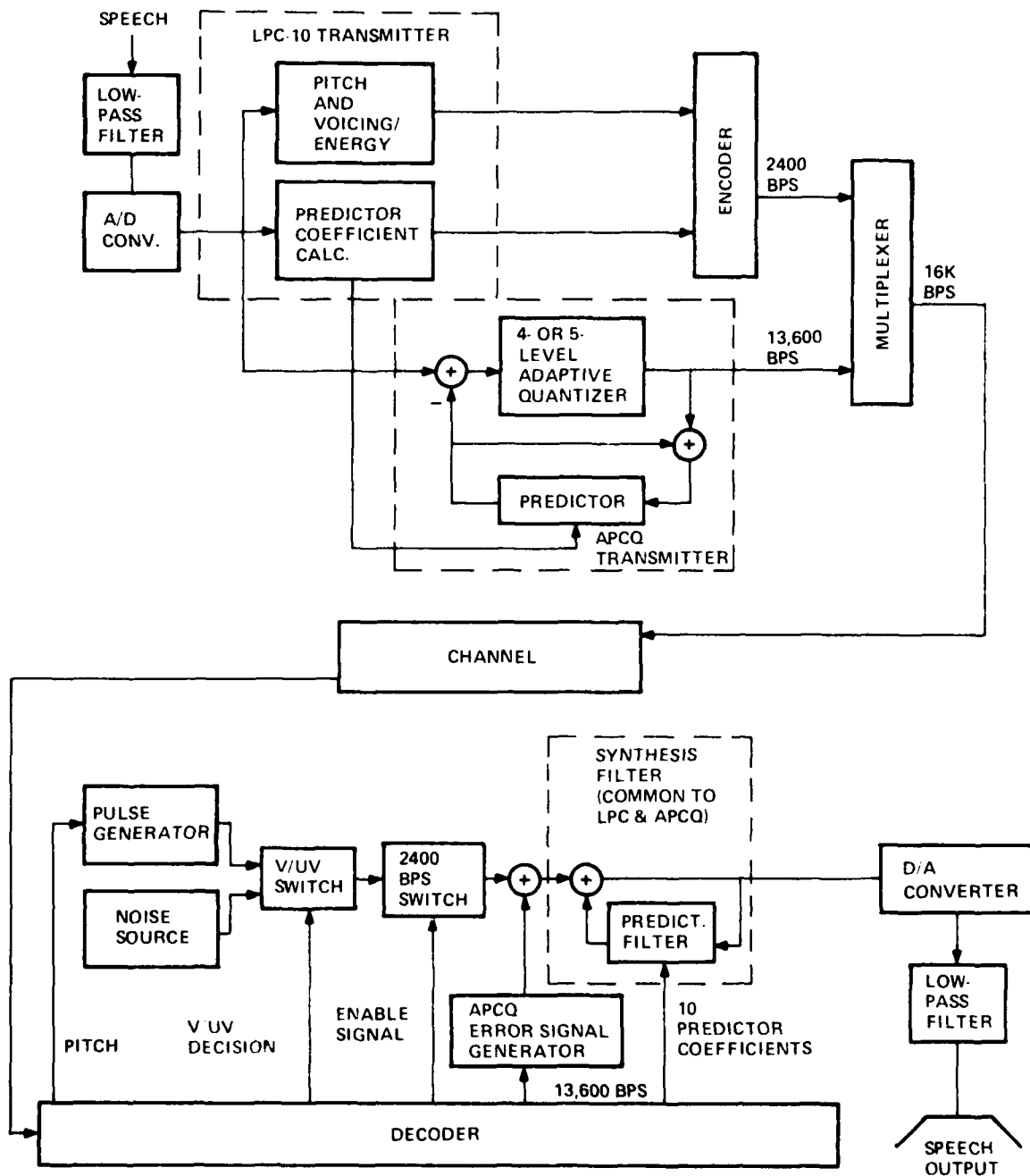
The embedded MRP scheme that utilizes 2.4 Kb/s LPC and 16.0 APCQ is shown in Figure 2-6.⁶ At the transmitter, the incoming speech is low-pass filtered and converted to digital by a PCM converter. A LPC algorithm calculates the reflection coefficients while separate logic determines pitch and voicing information. Many forms of LPC exist and any can be used for this operation.⁷ The output of the LPC analyzer is then encoded at 2400 bps. To obtain the remaining 13,600 bps, the input speech is then passed through an APCQ whose predictor uses the reflection coefficients or predictor coefficients as calculated in the LPC. Quantization of the residual signal is performed using either a 4- or 5-level adaptive quantizer as described by Jayant⁸ with modifications suggested by Goodman.⁹ After quantization, the residual signal is encoded to 13,600 bps. Since the 13,600 bps allocated to the residual signal is fixed, using more levels to encode each residual sample requires a lowering of the sampling rate. Table 2-1 indicates the input bandwidth for a 4- and 5-level quantizer, assuming that three samples, each having five levels, can be encoded in one 7-bit word.

Table 2-1 Input Bandwidth vs. Number of Residual Signal Quantizer Levels

Number of quantizing levels	Input bandwidth (Hz)
4	3400
5	2914

The choice of residual signal quantizer thus involves a tradeoff of input bandwidth versus processing noise. The 5-level quantizer introduces a little perceptible noise when heard through a telephone handset whereas the 4-level quantizer produces audible noise.

The operation of the receiver depends on whether it is receiving 2400 or 16,000 bps. Assuming it receives 16,000 bps of meaningful data, the synthesizer acts as a typical APCQ system. The 13,600 bps



8175 76E

FIGURE 2-6 INTEGRATED 2400/16,000 BPS LPC/APCQ VOICE DIGITIZER

are fed to the error signal generator which regenerates the residual signal. The residual signal enters the synthesis filter constructed from the received predictor coefficients and the output is converted to analog, filtered and presented to the listener.

Assuming instead that it receives only 2400 bps, the receiver then uses an LPC algorithm to reconstruct the voice. Here pitch information feeds a pulse generator, the voicing decision actuates a voiced/unvoiced switch, and an artificial residual signal excites the synthesis filter which is shared by the APCQ coder. Again the output is converted to analog, low pass filtered, and presented to the listener.

To indicate how this device eliminates synthesis of voice followed by analysis of voice at data rate changes, let us consider the following tandems: 16,000 to 2400 bps and 2400 to 16,000 bps. Figure 2-7 shows a transmitter operating at 16,000 bps. The frame rate is 22.5 ms implying that 360 bps are sent each frame. Fifty-four of these are for the LPC algorithm, divided up as shown in Figure 2-7. The remaining 306 bits are for the residual. At the data rate conversion point, the 306 bits of the residual signal are stripped off, leaving only the 54 bits for the LPC. The receiver reconstructs voice with this data.

Figure 2-8 illustrates the 2400 to 16,000 bps data rate change. the 54 bits/frame of LPC information is supplemented with 306 "dummy bits" which carry no information except, perhaps, to inform the receiver that they should be ignored. The 360 bits/frame (16,000 bps) travel over the 16,000-bps channel to the receiver which then uses only the 54 LPC bits to synthesize voice.

When the receiver has only the 2400 bps information, its voice quality and intelligibility is that of LPC. Voice quality is better than most channel vocoders but the voice still sounds synthetic and the pitch and voicing errors are audible and annoying.

When the receiver operates in the 16,000-bps mode, the received voice is natural and pitch and voicing errors do not

affect voice quality. Instead, the residual signal quantizer introduces speech related quantization noise which, for the 5-level quantizer, is barely perceptible when heard over the Western Electric U3 earpiece found in many telephone handsets. Though this 5-level quantizer produces higher quality speech than the 4-level one, it cannot be employed in the APCQ of the embedded coding scheme. This is due to the fact that the 5-level quantizer is only feasible if the input waveform has a bandwidth smaller than or equal to 2914 Hz. Unfortunately, at this bandwidth, the LPC algorithm does not perform well owing to the fact that accurate voicing decisions cannot be obtained in the absence of high frequency energy. Hence, to achieve reasonable performance in both the LPC and APCQ schemes, a compromise is to make use of a 4-level quantizer in encoding the error signal of APCQ thus making the overall system bandwidth 3400 Hz. In this configuration, unvoice/voice detections are relatively reliable for the LPC. Perceptually, the 4-level quantizer of the APCQ introduces a small amount of quantizing noise, but the resulting processed speech is still better than the 16,000 bps Continuously Variable Slope Delta Modulation (CVSD) based on our informal listening judgments.

The above LPC/APCQ system is capable of producing highly intelligible LPC encoded speech at 2.4 Kb/s and good quality APCQ processed speech at 16 Kb/s. By applying a 2-level quantizer on the error signal, the APCQ system is reduced to a modified APC that functions at 9.6 Kb/s. However, the resulting APC will perform suboptimally due to the omission of the pitch prediction loop. Often, tradeoffs between input bandwidths and quantizer levels have to be made in order to achieve satisfactory performance in all data rates. This illustrates a disadvantage inherent in all embedded MRP systems that the speech processing schemes are not independent of each other, and this often imposes a severe constraint on the design of these algorithms.

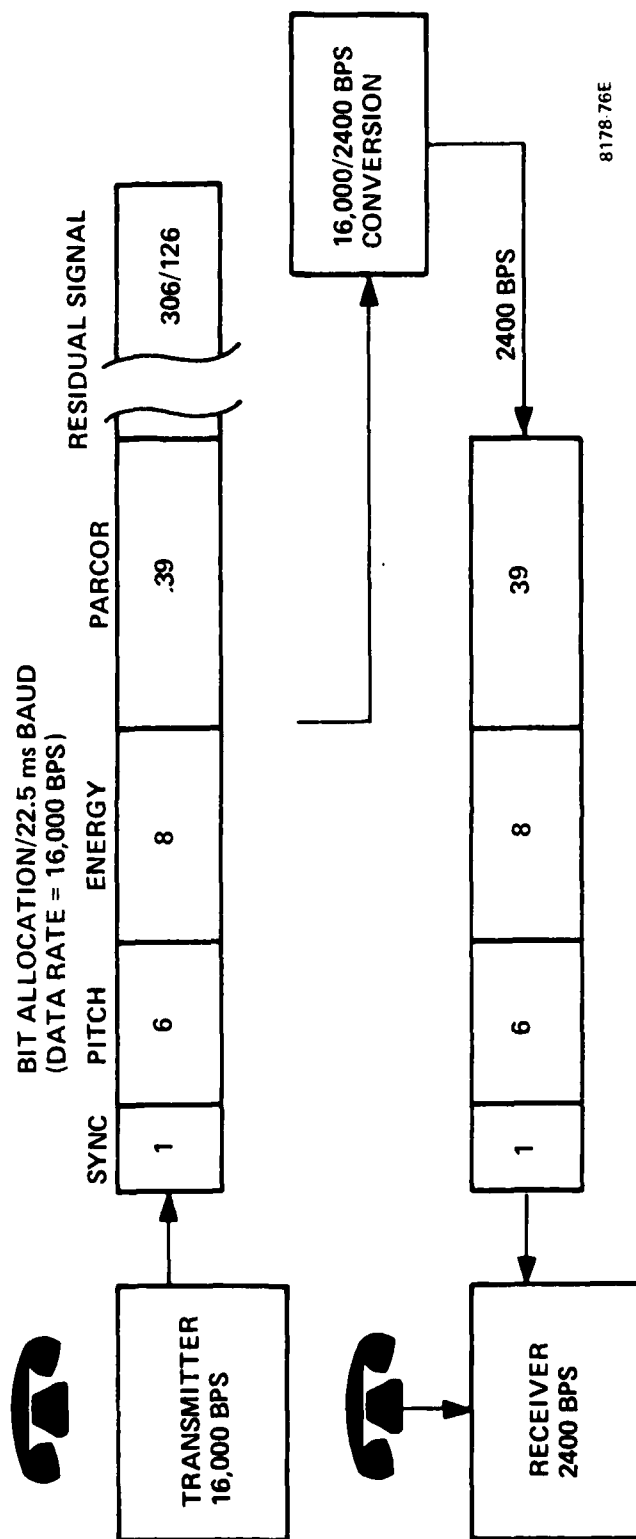


FIGURE 2-7 16,000 TO 2400 BPS TANDEM

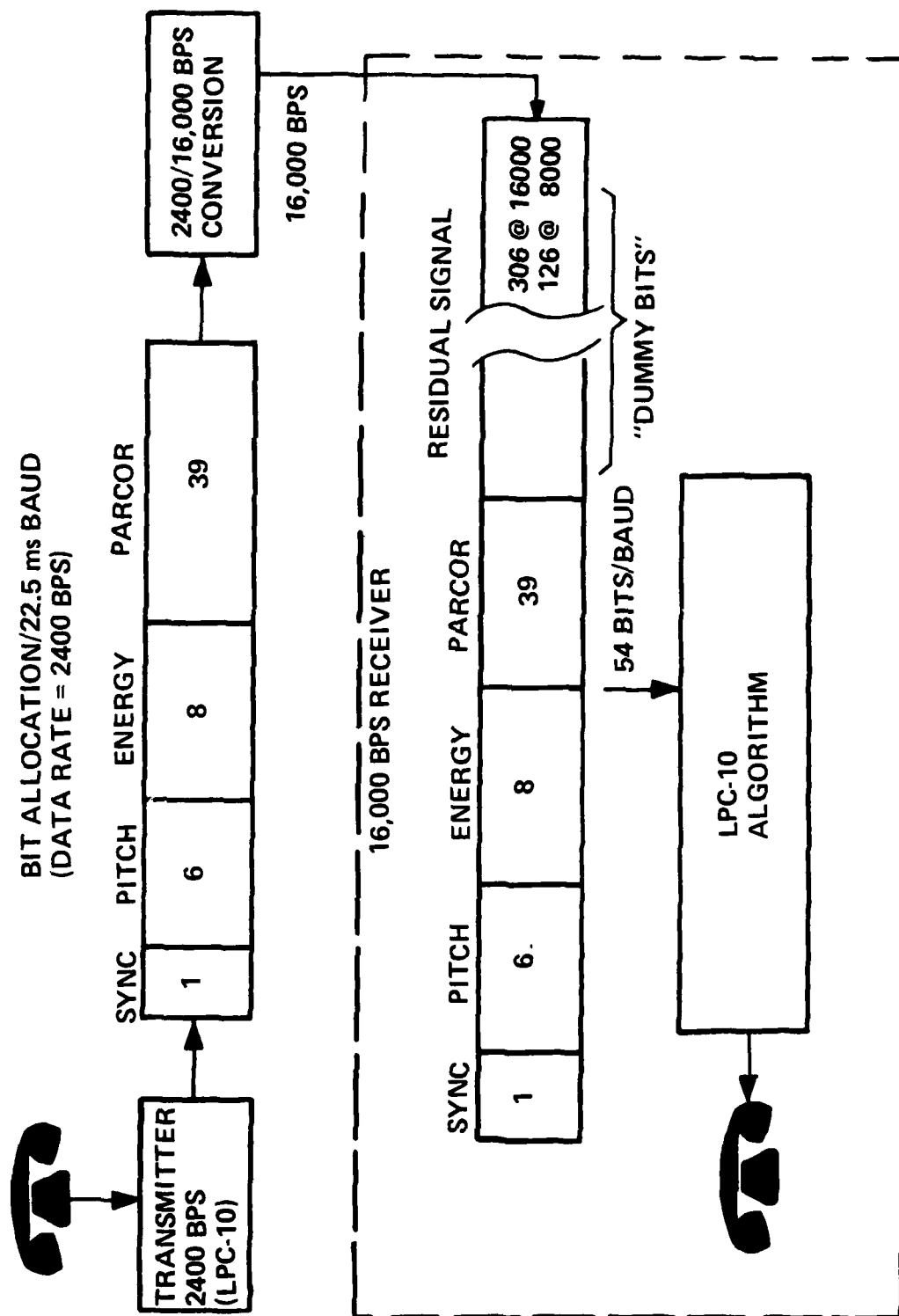


FIGURE 2-8 2400 TO 16,000 BPS TANDEM

8177-76E

2.4 The 2.4/8.0/9.6/16.0 Kb/s LPC/SBVC System

A second example of embedded MRP scheme is the 2.4/8.0/9.6/16.0 Kb/s Linear Predictive Coder/Split-Band Voice Coder.¹³ Block diagrams depicting the LPC/SBVC transmitter and receiver are shown in Figures 2-9 and 2-10, respectively. In this system, conventional LPC is used for the transmission of speech at 2.4 Kb/s, and employing a new technique known as Split-Band Voice Coding, the LPC residual signal is encoded which results in the higher transmission rates (8.0, 9.6, and 16.0 Kb/s). As shown in Figure 2-9, the transmitter first performs a linear prediction analysis on the incoming waveform which includes the computations of pitch, voicing decision, and predictor coefficients. Incorporating these parameters in two prediction loops, a residual signal of smaller dynamic range is generated in the same manner as shown in conventional APC schemes. Then a split-band technique is applied to partition the frequency band of the error signal into subbands, each of which is individually quantized. In particular, the method calls for the use of a 3-stage tree structure of quadrature mirror filters (QMF) to split the error signal band into 8 subbands.¹⁴ For an input signal of 4000 Hz bandwidth, subbands of 500 Hz are resulted. At the first stage of the transmitter, the QMF filters split the input into two bands; that is, 0-2000 Hz and 2000-4000 Hz. Then a downsampling procedure is utilized to reduce the number of samples by a half. At the second stage, the identical bandsplitting process is applied to each subband. Consequently, 2 more subbands are generated each of which has a 1000 Hz bandwidth. At the end of the second stage, a total of 4 subbands is obtained. The method is repeated one more time and eight subbands, each of which is 500 Hz wide spanning frequency from 0 to 4000 Hz, are created. After performing the bandsplitting process, the subband signals are individually quantized for transmission. As it is pointed out in Section 2.2.2, the embedded MRP scheme utilizes only one speech processing algorithm (e.g., LPC/SBVC), but the utilization of different level quantizers in encoding the output parameters results in the transmission of several data rates. In the case of SBVC, by encoding the subband signals with quantizers of different bits, transmissions at 8.0, 9.6, and 16.0 data rates are possible. Bit allocations for the LPC/SBVC embedded

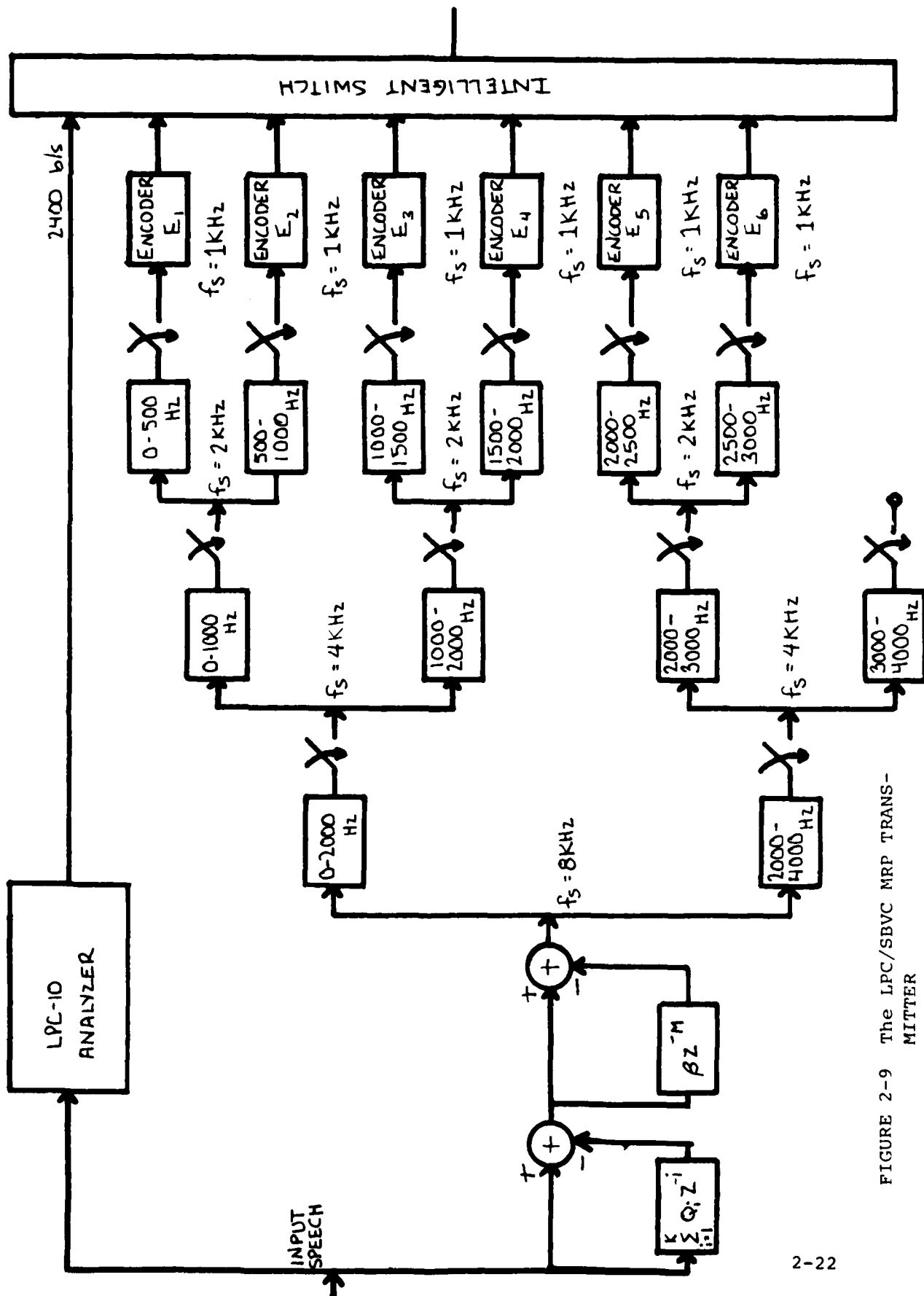


FIGURE 2-9 The LPC/SBVC MRP TRANSMITTER

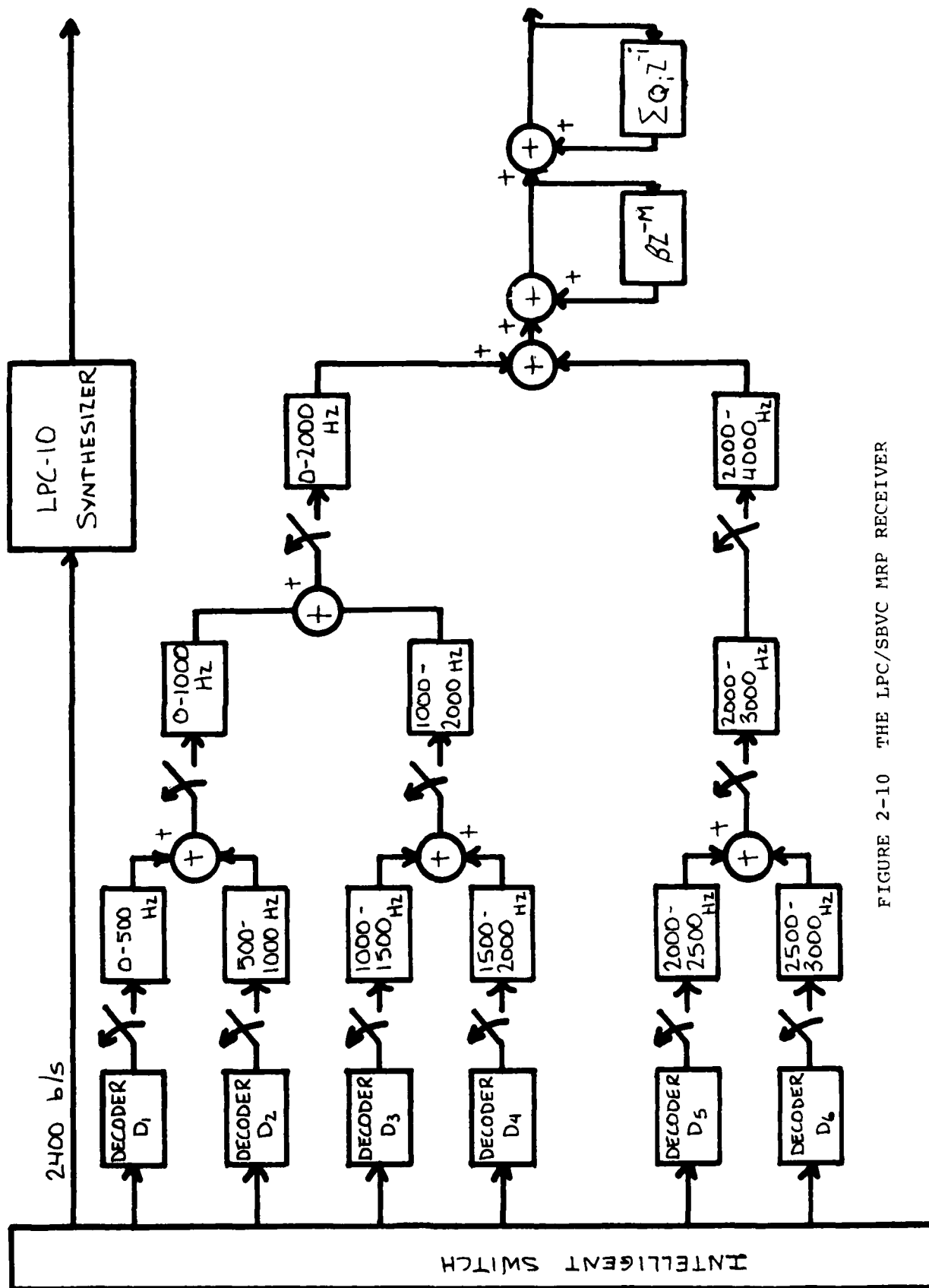


FIGURE 2-10 THE LPC/SBVC MRP RECEIVER

MRP scheme for the various rates are shown in Table 2-2.

Depending on the data rate, different synthesizers are used to reconstruct the input waveform. For the 2.4 Kb/s data rate, a conventional LPC synthesizer is utilized. As for the higher data rates, reverse operations of the SBVC transmitter, which include up-sampling and bandpass filtering, are performed. The resulting subband signals are later recombined yielding an estimate of the LPC residual signal. Then the original signal is reconstructed using an APC synthesizer as shown in Figure 2-10.

To illustrate the operations of the LPC/SBVC system, transmitter output bits are considered. For the 2.4 Kb/s rate, a frame of 180 input samples are brought in every 22.5 msec where a tenth order linear prediction analysis is performed resulting in 54 output bits per frame. For the higher rates, additional bits are utilized to characterize the error signal. As in the case of 8.0 Kb/s, an extra 128 bits are employed yielding a total of 180 bits of frame. For 9.6 Kb/s, the output frame length is 216 bits whereas 360 bits per frame are outputted from the 16.0 Kb/s transmitter. Furthermore, to understand how the LPC/SBVC system eliminates synthesis of voice followed by analysis of voice at data rate changes, the tandems between all possible data rates as indicated in Figure 2-11 have to be considered. For the sake of simplicity, let us only examine, in detail, the 2400/16,000 bps tandem.

To convert the 16 Kbps data rate into 2.4 Kbps, 306 bits used to encode the splitband filtered LPC error signal are discarded leaving only the 54 LPC parameter bits which are transmitted to the 2.4 Kbps synthesizer. On the other hand, to change the 2.4 Kbps to 16 Kbps, 306 zero or "dummy" bits have to be inserted at the switch or data rate conversion point and the entire 360 bits are passed to the 16 Kbps receiver. Realizing only 54 bits out of the 360 received are of importance, an LPC synthesizer is then utilized to reconstruct the input waveform. Employing these LPC/SBVC terminals in conjunction with intelligent switches, communications between wideband and narrowband terminal users are made possible.

DATA RATE	# OF BITS/ SEC NEEDED TO ENCODE LPC PARAMETERS	ENCODER E ₁		ENCODER E ₂		ENCODER E ₃		ENCODER E ₄		ENCODER E ₅		ENCODER E ₆		OVER- HEAD RATE IN BITS/ SEC
		# of Quan- tizer Levels	total # bits/ sec	# of Quan- tizer Levels	total # bits/ sec	# of Quan- tizer Levels	total # bits/ sec	# of Quan- tizer Levels	total # bits/ sec	# of Quan- tizer Levels	total # bits/ sec	# of Quan- tizer Levels	total # bits/ sec	
2.4 Kb/s	2.4 K													
8.0 Kb/s	2.4 K	2	1 K	2	1 K	2	1 K	2	1 K	2	1 K			.6K
9.6 Kb/s	2.4 K	4	2 K	2	1 K	2	1 K	2	1 K	2	1 K			1.2K
16.0 Kb/s	2.4 K	8	3 K	8	3 K	4	2 K	4	2 K	2	1 K	2	1 K	1.6K

TABLE 2-2 BIT ALLOCATIONS FOR 2.4/8.0/9.6/16.0 KB/S LPC/SEVC SCHEME

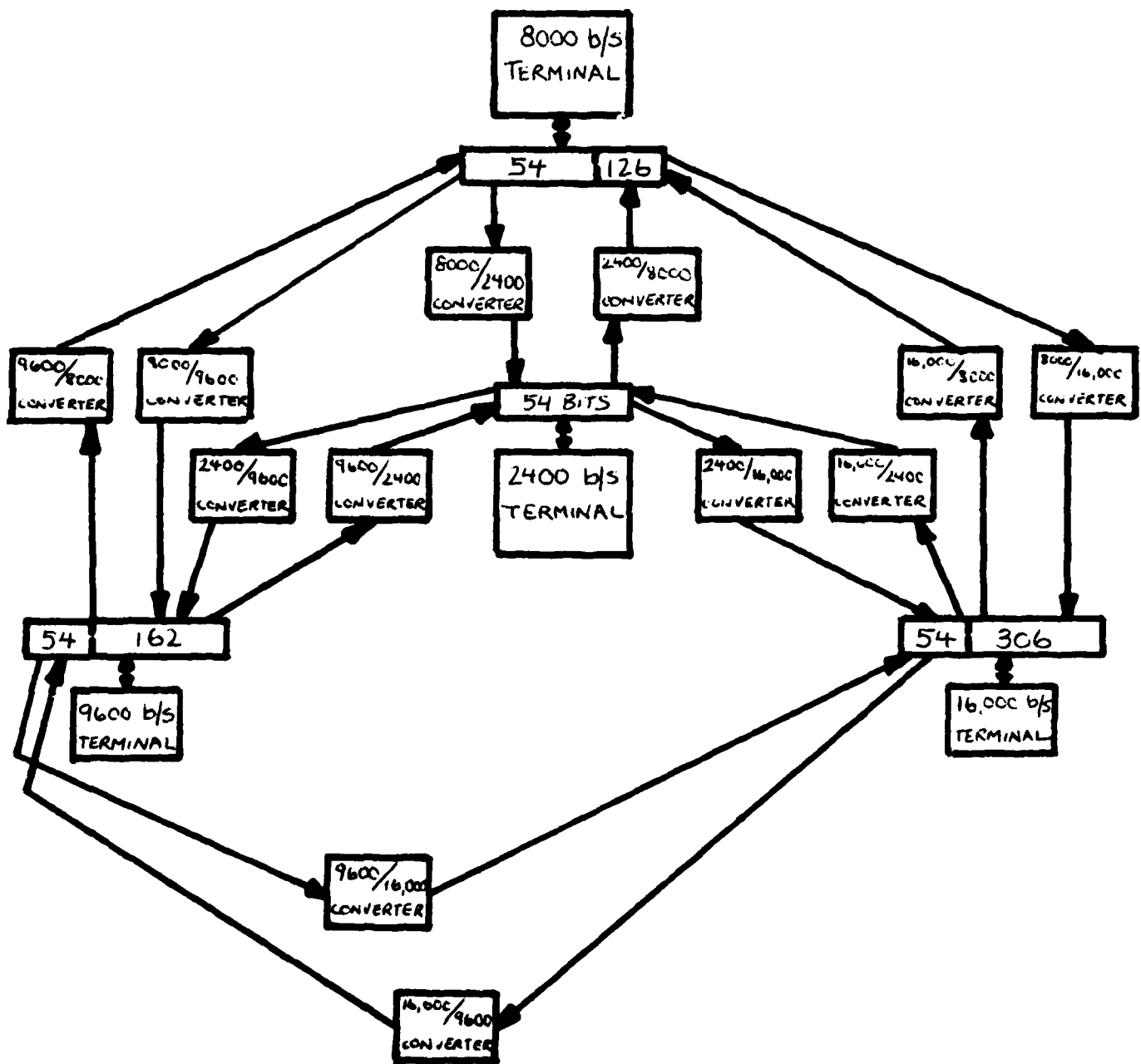
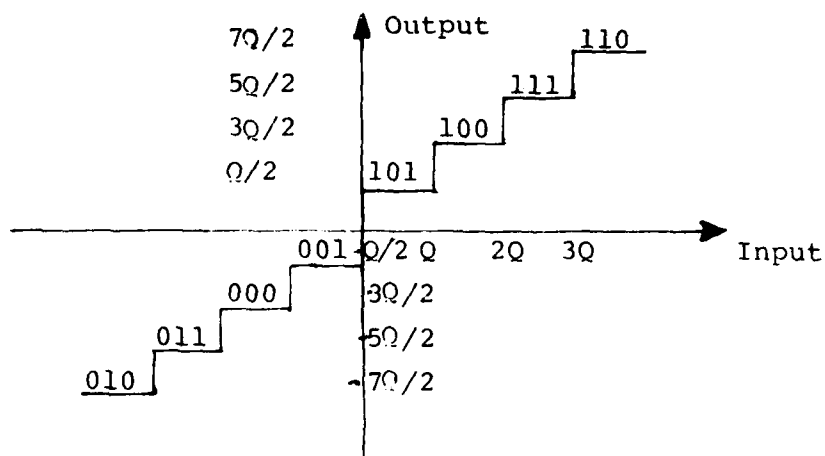


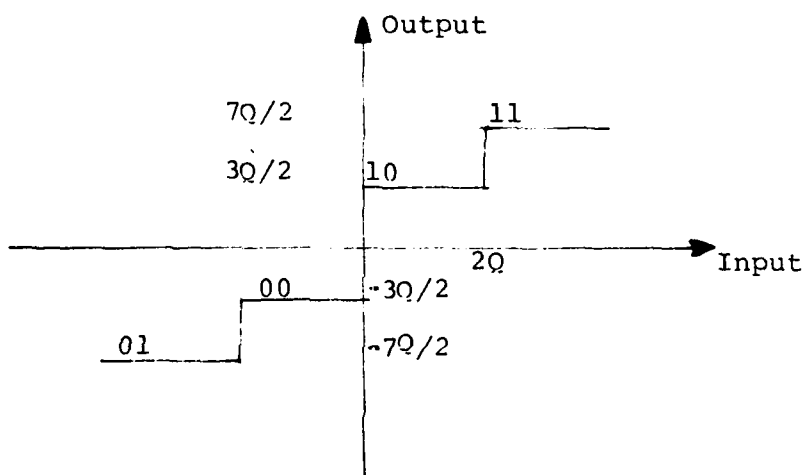
FIGURE 2-11 2.4, 8.0, 9.6, 16.0 KBPS TANDEM CONFIGURATIONS

Operations needed to change data rates between 2.4 K and 16.0 Kb/s in the LPC/SBVC system are identical to that of the LPC/APCQ. However, conversions between any two higher rates call for the use of embedded coding. To illustrate this, let us consider subband #1 in the LPC/SBVC scheme. For the 16 Kb/s data rate, a 3-bit quantizer, as shown in Figure 2-12(a), is utilized to quantize the subband waveform. For the 9.6 Kb/s data rate, only a 2-bit one, as depicted in Figure 2-12(b), is employed. As for the 8.0 Kb/s, a 1-bit quantizer shown in Figure 2-12(c) is applicable. Hence, conversion from a high data rate to a low one requires the intelligent switch to strip out the correct bits and transmit them to the receiver. For example, to reduce the data rate from 16 Kb/s to 9.6 Kb/s, the switch has to derive two bits from the available three for each sample of subband 1. If code words of the 3-bit quantizer are chosen as shown in Figure 2-12(a), then the switch will only retain the first two bits resulting in a 2-bit quantizer whose code words are shown in Figure 2-12(b). Similarly, conversion from 16 Kb/s to the 8 Kb/s is achieved if only the first bit is kept for every 3-bit code word. At the receiver, a de-quantization procedure is performed which converts the code words back to output levels.

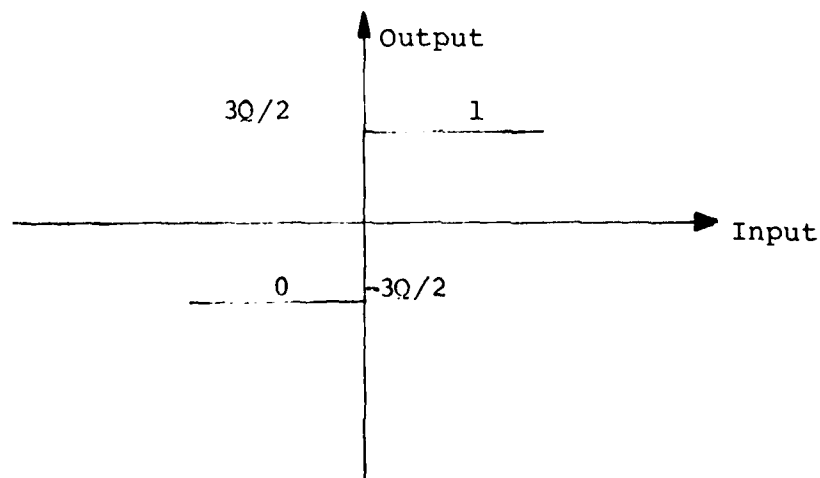
On the other hand, to convert data rates from a lower rate to a higher one, "dummy" or zero bits are inserted at the end of each output code word of the lower rate. To illustrate this, let us consider subband #1 of the LPC/SBVC system. Conversion from the 8.0 Kb/s encoder to the 9.6 Kb/s one only requires the insertion of a zero for every 1-bit code word outputted from the lower data rate scheme. In the same manner, additions of two zero bits for every code word outputted will boost the data rate from 8.0 to 16.0 Kb/s. Hence, utilizing this embedded coding procedure, quantizations for subband signals can be converted easily from one data rate to another. Furthermore, the similar strategy has to be applied to the encoding of the overhead bits in order to make the LPC/SBVC algorithm to be truly embedded.



(a) a 3-bit Quantizer



(b) a 2-bit Quantizer



(c) a 1-bit Quantizer

FIGURE 2-12 1, 2, and 3-BIT QUANTIZERS WITH EMBEDDED CODING

2.4.1 The Theory of Quadrature Mirror Filter

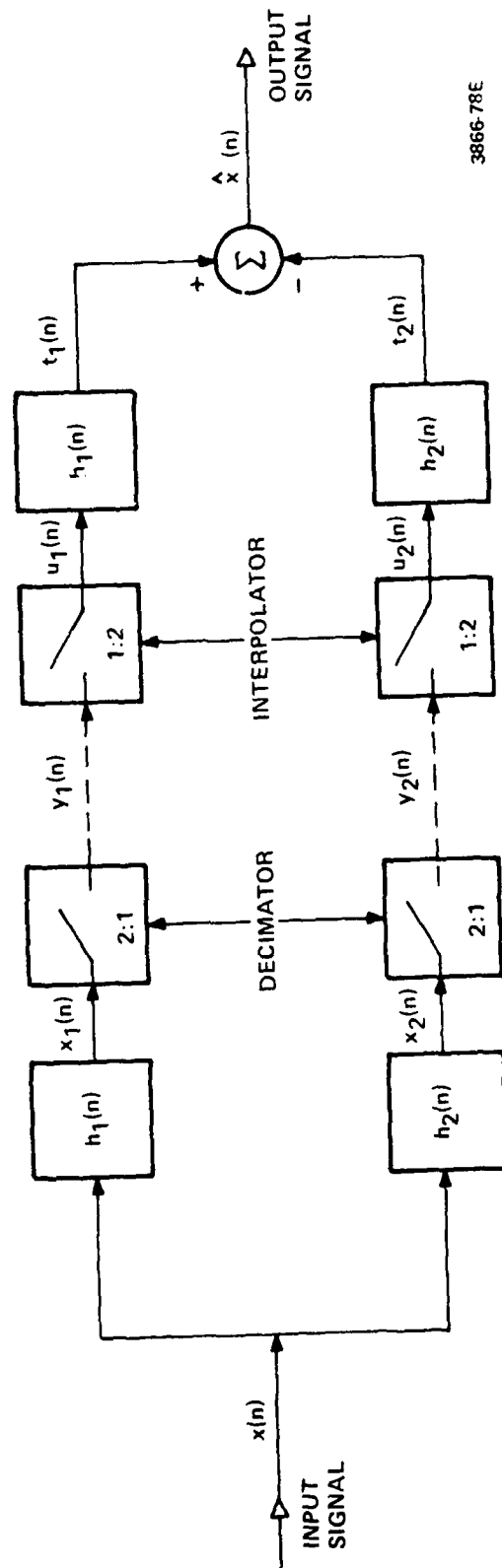
In the Split Band Voice Coding systems, quadrature mirror filters (QMF) are utilized to band-split/reconstruct the input waveform via decimation/interpolation methods¹¹. This section shows that the use of QMF filters will achieve perfect splitting/reconstruction without any spectral aliasing.

For explanatory purposes, consider the ideal splitting/reconstruction process described in Figure 2-13. For this system, the following definitions apply:

- a. $x(n)$ is a Nyquist band-limited residual signal with z-transform $X(z)$.
- b. h_1 is the impulse response of the low-pass filter and the z-transform of which is $H_1(z)$.
- c. $h_2(n)$ is the impulse response of the high-pass filter and the z-transform of which is $H_2(z)$.
- d. $y_1(n)$ is a baseband equivalent low-pass signal with z-transform $Y_1(z)$.
- e. $y_2(n)$ is a baseband equivalent high-pass signal with z-transform $Y_2(z)$.

The signal $x(n)$, is processed by filters $h_1(n)$ and $h_2(n)$ yielding the low-pass and high-pass equivalents, $x_1(n)$ and $x_2(n)$, of the residual signal. As their spectra occupy half the Nyquist bandwidth of the original signal, the sampling rate in each band can be halved by decimating (ignoring) every second sample. For reconstruction, the signals $y_1(n)$ and $y_2(n)$ are interpolated by inserting one zero-valued sample every other time and then filtered, respectively, by $h_1(n)$ and $h_2(n)$ before being added to give the signal $\hat{x}(n)$. The dashed lines, shown in Figure 2-8, represent the data passed to the communication channel(s) by the speech processing system.

FIGURE 2-13 BAND-SPLITTING AND RECONSTRUCTION USING SINGLE-STAGE QMF



3866-78E

In order to minimize $(x(n) - \hat{x}(n))$, certain restrictions on the filters, $h_1(n)$ and $h_2(n)$, must be met. We will derive these restrictions by constructing the transfer function of the QMF structure.

Using z-transform notation and referring to Figure 2-11, we may write the intermediary filtered output as

$$X_1(z) = H_1(z) X(z) \quad (2-1)$$

and

$$X_2(z) = H_2(z) X(z) \quad (2-2)$$

The transforms of the decimated signals, $y_1(n)$ and $y_2(n)$, and of the interpolated signals, $u_1(n)$ and $u_2(n)$, are given by:

$$Y_1(z) = 1/2 (X_1(\tilde{z}) + (X_1(-\tilde{z}))), \quad \tilde{z} = z^{1/2} \quad (2-3)$$

$$Y_2(z) = 1/2 (X_2(\tilde{z}) + (X_2(-\tilde{z}))) \quad (2-4)$$

$$U_1(z) = Y_1(z^2) \quad (2-5)$$

$$U_2(z) = Y_2(z^2) \quad (2-6)$$

After the final filtering operating, the transforms of the reconstructed waveform components, $t_1(n)$ and $t_2(n)$, are given by

$$T_1(z) = H_1(z) U_1(z) \quad (2-7)$$

$$T_2(z) = H_2(z) U_2(z) \quad (2-8)$$

Using the relations expressed in (2-3) through (2-8), the z-transforms can be rewritten as

$$T_1(z) = 1/2 (H_1(z) X(z) + H_1(-z) X(-z)) H_1(z) \quad (2-9)$$

$$T_2(z) = 1/2 (H_2(z) X(z) + H_2(-z) X(-z)) H_2(z) \quad (2-10)$$

The z-transform of the reconstructed waveform, $\hat{x}(n)$ is obtained by adding (2-9) and (2-10)

$$\hat{X}(z) = 1/2 (H_1^2(z) - H_1^2(-z)) X(z) + 1/2 (H_1(-z) H_1(z) - H_2(-z) H_2(z)) X(-z) \quad (2-11)$$

If we assume that

$$H_2(z) = H_1(-z) \quad (2-12)$$

then the reconstructed waveform transform becomes

$$\hat{X}(z) = 1/2 (H_1^2(z) - H_1^2(-z)) X(z) \quad (2-13)$$

Evaluating z on the unit circle gives the Fourier transform of $\hat{X}(z)$

$$\hat{X}(e^{j\omega T}) = 1/2 (H_1^2(e^{j\omega T}) + H_1^2(e^{j(\omega + \frac{\omega_s}{2})T})) X(e^{j\omega T}) \quad (2-14)$$

For the case when $h_1(n)$ is an even, symmetrical FIR filter of order N , then it can be shown that (2-14) reduces to

$$\hat{X}(e^{j\omega T}) = 1/2 e^{-j(N-1)\omega T} X(e^{j\omega T}) \quad (2-15)$$

where $H_1^2(e^{j\omega T})$ exhibits an odd symmetric property about $\omega_s/4$ and the half-power point $H_1^2(\omega_s/4) = 0.5$.

The inverse transform yields a perfectly reconstructed signal (no frequency distortion) with a gain factor of $1/2$ and delay of $N-1$ samples as shown by

$$\hat{X}(n) = 1/2 x(n - N + 1) \quad (2-16)$$

Therefore, we have shown that to guarantee perfect reconstruction of the original LPC residual spectrum, the following filter constraints must be satisfied

$$h_1(n): \text{symmetrical, even order FIR} \quad (2-17)$$

$$H_2(z) = H_1(-z) \text{ or } h_2(n) = (-1)^n h_1(n); n=0,1,\dots,N-1 \quad (2-18)$$

$$H_1^2(z) + H_2^2(z) = 1 \quad (2-19)$$

Throughout the formulation of the bandsplitting/reconstruction process with QMF filters, there is no stipulation on the length of the FIR filter (as long as it is even). Hence, perfect splitting/reconstruction can be achieved with relatively short filters. An example of QMF filters is given by the 12-tap one whose coefficients are tabulated in Table 2-3.¹² Frequency response of the filter is depicted in Figure 2-14. As illustrated in the Figure, the filter is characterized by a flat-passband response, a 3 dB point at 2000 Hz, and relatively small stopband rejection. However, the composite frequency response after 3 stages of bandsplitting and reconstruction yields only 0.5 dB of ripple as shown in Figure 2-15. Therefore, this 12-tap filter is employed in the study and real-time implementation of the LPC/SBVC scheme.

Low-Pass Filter Coefficients	High-Pass Filter Coefficients
$h_1(1) = -3.$	$h_2(1) = -3.$
$h_1(2) = 10.$	$h_2(2) = -10.$
$h_1(3) = -4.$	$h_2(3) = -4.$
$h_1(4) = -24.$	$h_2(4) = 24.$
$h_1(5) = 28.$	$h_2(5) = 28.$
$h_1(6) = 120.$	$h_2(6) = -120.$
$h_1(7) = 120.$	$h_2(7) = 120.$
$h_1(8) = 28.$	$h_2(8) = -28.$
$h_1(9) = -24.$	$h_2(9) = -24.$
$h_1(10) = -4.$	$h_2(10) = 4.$
$h_1(11) = 10.$	$h_2(11) = 10.$
$h_1(12) = -3.$	$h_2(12) = 3.$

TABLE 2-3 TABULATION OF COEFFICIENTS FOR BOTH THE
LOW-PASS AND HIGH-PASS QMF FILTERS

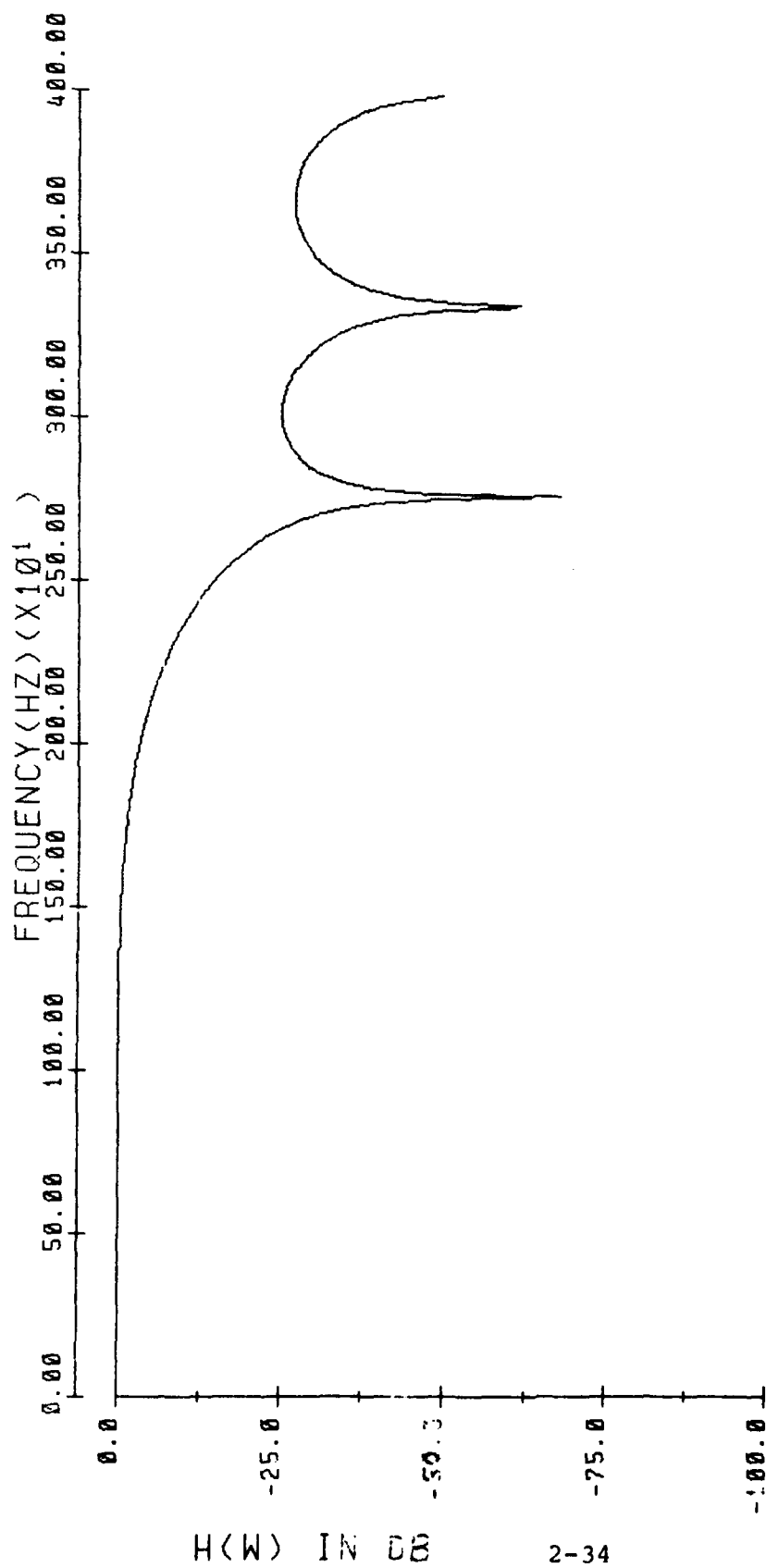


FIGURE 2-14 MAGNITUDE RESPONSE OF THE 12-TAP QMF FILTER

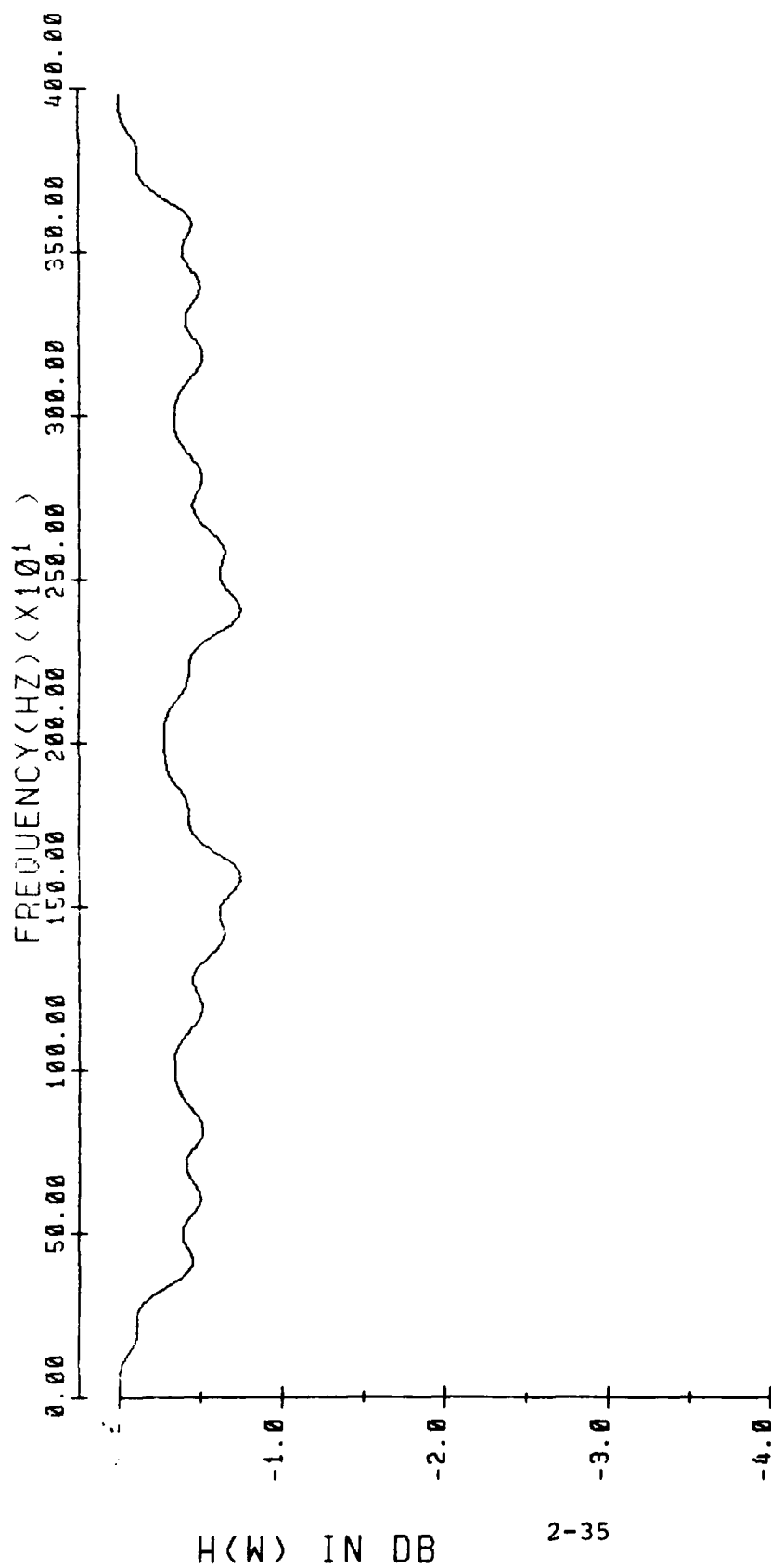


FIGURE 2-15 COMPOSITE MAGNITUDE RESPONSE OF A
SPLIT-BAND VOICE CODER(SBVC)
USING 3 STAGES OF QMF FILTERS
(NO QUANTIZATIONS)

2.4.2 Results of the LPC/SBVC System

In contrast to the LPC/APCQ system as discussed in Section 2.3, the LPC/SBVC is more versatile since it can be utilized to transmit speech at 2.4, 8.0, 9.6, and 16.0 Kb/s. This can be attributed to the fact that the latter scheme does not have a rigid relationship between the full input signal bandwidth and quantizer levels. Indeed, by splitting the band of the LPC residual signal into subbands and encoding each one separately, the LPC/SBVC system trades part of the output speech bandwidth off with data rate. To further illustrate this, the bandwidth of the 8.0 or 9.6 Kb/s system, as shown in Table 2-2, is only 2500 Hz whereas the 16.0 Kb/s system has a 3 KHz bandwidth. Unfortunately, in comparison to conventional APC or APCQ at 8.0, 9.6, or 16.0 Kb/s, the LPC/SBVC scheme does not produce good quality speech outputs. This can be partly explained by the fact that splitband voice coding is not directly applied to the input signal. Instead, it is utilized in the coding of the LPC error signal which spectrally is flatter than the original one. In this situation, split-band voice coding system may not be as advantageous. Furthermore, as a result of the split-band filters, quantizers cannot be configured within the prediction loop. Accumulations of quantizing errors greatly hamper the success of such coders and this result has been substantiated by recent reports^{12,15}

SECTION III

Conclusions and Recommendations

3.1 Conclusions

This contract has resulted in i) the development of two embedded multiple-rate processing (MRP) schemes, namely, the 2.4/16.0 Kb/s Linear Predictive Coding/Adaptive Predictive Coding with Adaptive Quantization (LPC/APCQ) and the 2.4/8.0/9.6/16.0 Kb/s Linear Predictive Coding/Split Band Voice Coding (LPC/SBVC); ii) the real-time implementation of the LPC/SBVC algorithm on the Sylvania Programmable Signal Processors (PSP). Both schemes utilize LPC as the 2.4 Kb/s coder since LPC is known to produce highly intelligible speech at 2.4 Kb/s data rate. For the LPC/APCQ scheme, APCQ is employed to encode the LPC residual signal and this results in good quality speech at 16 Kb/s which is relatively insensitive to pitch and voicing mistakes. Unfortunately, this method does not perform well in the medium-band transmission (8-10 Kb/s) owing to the strict relationship between the full input bandwidth and the levels of error signal quantizers. On the other hand, the LPC/SBVC is more versatile in the sense that it functions at 2.4, 8.0, 9.6, and 16.0 Kb/s. In contrast to the LPC/APCQ algorithm, LPC/SBVC employs split-band technique to encode the LPC residual signal. By partitioning the full bandwidth of the input into subbands, each of them is quantized differently to obtain the various data rate. Unlike the LPC/APCQ, there exists no direct relationship in the LPC/SBVC scheme between the full bandwidth of the input signal and the quantizer levels. Instead, the latter method trades off the number of quantized subbands and data rates. Unfortunately, when compared to conventional APC or APCQ schemes, the LPC/SBVC system does not produce good quality speech at the high data rates. This is probably due to the configuration of the SBVC quantizers which are outside of the APC predictor loops.

3.2 Recommendations

In this study, the concept of embedded multiple-rate processing schemes and their utility in facilitating narrowband/wideband communications are presented. This idea of being able to use a single "universal" voice digitization algorithm to encode speech for a variety of data rates is indeed very appealing and it should be pursued further.

Though the LPC/SBVC algorithm discussed in this report does not provide the speech quality as good as expected at the higher data rates, however, it illustrates the fact that encoding schemes in the frequency domain are the most flexible in achieving a list of different data rates. One of the reasons is that in the frequency domain, reduction in allocation of bits does not affect the entire frequency band. Instead, distortions are only localized in a particular frequency region which may not be perceptually noticeable. Recently, new frequency domain speech processing techniques, such as adaptive transform coding (ATC), have been studied and they are known to produce high quality processed outputs at data rates above 8 Kb/s.^{13,14} Since these ATC algorithms have great potential in embedded MRP applications, they should be further investigated.

REFERENCES

1. B. S. Atal and S. L. Hanauer, "Speech Analysis and Synthesis by Linear Prediction of the Speech Wave," JASA Vol. 50, No. 2, pp. 637-655, 1971.
2. R. F. Steele, "Delta Modulation Systems," John Wiley and Sons, New York, 1975.
3. M. R. Sambur and N. S. Jayant, "LPC Analysis/Synthesis from Speech Inputs Containing Quantization Noise or Additive White Noise," IEEE Trans. Acoustics, Speech and Signal Processing, Vol. ASSP-24, No. 6, 1976.
4. A. J. Goldberg, and H. L. Shaffer, "A Real-Time Adaptive Predictive Coder Using Small Computers," IEEE Trans. Communication, Vol. COM-23, No. 12, pp. 1443-1451, December 1975.
5. A. J. Goldberg, "2400/16,000 BPS Multirate Voice Processor," IEEE International Conference on Acoustics, Speech and Signal Processing, Tulsa, Oklahoma, 1978.
6. A. J. Goldberg, R. L. Freudberg, and R. S. Cheung, "High Quality 16 Kb/s Voice Transmission," IEEE International Conference on Acoustics, Speech and Signal Processing, Philadelphia, PA, 1976.
7. J. D. Markel and A. H. Grey, "A Linear Predictive Vocoder Simulation Based on the Autocorrelation Method," IEEE Trans. Acoust. Speech and Signal Processing, Vol. ASSP-22, pp. 124-134, April 1974.
8. N. J. Jayant, "Digital Coding of Speech Waveforms: PCM, DPCM, and DM Quantizers," Proc. IEEE, Vol. 62, No. 5, pp. 611-632, May 1974.
9. D. L. Goodman and R. M. Wilkinson, "A Robust Adaptive Quantizer," IEEE Trans. Communications, Vol. COM-23, No. 11, pp. 1312-1365, November 1974.

10. D. Esteban and C. Galand, "Application of Quadrature Mirror Filters to Split Band Voice Coding Schemes," IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 191-195, Hartford, Conn., 1977.
11. A. Croisier, D. Esteban, and C. Galand, "Perfect Channel Splitting by use of Interpolation/Decimation/Tree Decomposition Techniques," 1976 International Conference on Information Sciences and Systems, Patras.
12. Final Report, "APC-4 Speech Improvement Study," Contract No. BOA MDA 904-77-A0-09915, July 1978.
13. R. Zelinski and P. Noll, "Adaptive Transform Coding of Speech Signals," IEEE Trans. Acoustics, Speech and Signal Processing, Vol. ASSP-25, No. 4, pp. 299-309, August 1977.
14. J. M. Tribolet and R. E. Crochiere, "A Vocoder Driven Adaptation Strategy for Low-Bit Rate Adaptive Transform Coding of Speech," Proc. 1978 Int. Conf. on Digital Signal Processing, Florence, Italy, September 1978.
15. J. Makhoul and M. Berouti, "Adaptive Noise Spectral Shaping and Entrophy Coding in Predictive Coding of Speech," IEEE Trans. Acoustic, Speech and Signal Processing, Vol ASSP-27, No. 1, February 1979.

